

УДК 004.7

Анализ состояния растительности и оценка рисков развития сельскохозяйственных культур на основе статистической теории обучения

Ю.А. Грипич

Институт космических исследований НАНУ-НКАУ
inform@ikd.kiev.ua

Abstract

Grypich. Yu.,. Analysis of Crops State and Risk Assessment Based on Statistical Learning Theory. This paper solves the task of agricultural monitoring and risk assessment with the respect to the crops development and droughts. It is solved using different information (satellite and ground and simulation data). Mathematical apparatus used for solving this problem is being asserted. The results of its practical usage for crops development risk assessment are proposed.

Введение

Сельское хозяйство является одним из наиболее приоритетных направлений развития экономики Украины. При этом наша страна страдает от глобальных изменений климата, стихийных бедствий и засух, низких температур, наводнений, перепадов температуры. В последние годы такая тенденция стала еще более угрожающей. Например, весной и летом 2007 года южные регионы Украины пострадали от существенной засухи. По данным Министерства Чрезвычайных ситуаций Украины вследствие засухи были полностью уничтожены посевы площадью около 1400 тыс.га (16% засеянных площадей) и были массово повреждены посевы сельскохозяйственных культур общей площадью около 8500 тыс. Га. Общая сумма ущерба – свыше 450 млн. гривен (чрезвычайная ситуация государственного уровня). Страны ЕС тоже регулярно потерпают от сильных засух. Общая сумма ущерба от засух за последние 30 лет составила около 100000 млн. евро.

Поэтому органам государственной власти всех уровней очень важно иметь средства мониторинга состояния сельскохозяйственных культур и оценки рисков их развития, таким образом, иметь возможность решения актуальных заданий государственного значения, а именно прогнозирование урожайности, оперативной оценки площадей посевов, прогнозирования валовых сборов сельскохозяйственной продукции, задач продовольственной безопасности и т.д.. Но сейчас в Украине системы и средства мониторинга практически полностью отсутствуют.

В мире активно развиваются комплексные

системы мониторинга состояния растительности и оценки рисков. Так, в Соединенных Штатах Америки разработкой таких систем активно занимаются специалисты агентства NOAA (<http://www.nws.noaa.gov/>), в Европе используется система мониторинга MARS, в рамках Европейской программы GMES выполняется проект Geoland/Geoland2 (<http://www.gmes.info/pages-principales/projects/land-projects/geoland2/>), в Российской Федерации аналогичная система мониторинга разработана специалистами Института космических исследований РАН.

Важная задача мониторинга растительности и оценки рисков отмечена также в рабочих планах межправительственной группы GEO (Group on Earth Observations) в разделе «AG-07-03: Global Agricultural Monitoring» (создание глобальной системы мониторинга сельского хозяйства, управления рисками в аграрном секторе), а также является одним из важнейших прикладных направлений исследований Объединенного Исследовательского центра Европейской комиссии. Также стоит отметить, что вопросами мониторинга сельскохозяйственных культур активно занимаются участники международной организации по вопросам продовольствия и сельского хозяйства ООН (Food and Agriculture Organization of the United Nations).

Для мониторинга сельскохозяйственных культур и оценки рисков в мире используют геопространственные данные разной природы: аэрокосмические снимки и продукты (например, цифровую модель рельефа, карты землепользования), а также данные двумерного или трехмерного моделирования (в частности, метеорологические или физические модели). Результатом мониторинга при использовании такого рода информации становятся цифровые

карты или многослойные геопространственные данные, существенно облегчающие процесс принятия решения соответствующими органами. Подобная информация может использоваться не только для картографирования районов с повышенным риском во время или после неблагоприятного события (засухи, заморозков, града), но и на других этапах цикла развития стихийного бедствия (disaster cycle) — в том числе для построения карт рисков, иллюстрирующих вероятность наступления события и ущерб, который может быть им причинен.

В данной работе решается задача сельскохозяйственного мониторинга и оценки рисков, связанных с развитием сельскохозяйственных культур и засухами, на основе разнородной информации (спутниковых и наземных данных, а также данных моделирования), обосновывается применяемый для решения этой задачи математический аппарат, а также предлагаются результаты его практического применения для оценки рисков развития сельскохозяйственных культур в Украине.

Постановка задачи оценки сельскохозяйственных рисков на основе геопространственной информации и метод ее решения

В теории статистического обучения ожидаемая величина потерь определяется функционалом риска (risk functional) [1]

$$R(w) = \int L(d, F(x, w)) dF_{x,d}(x, d)$$

где $L(d, F(x, w))$ — функция потерь, а интеграл берется по всем возможным значениям (x, d) . Целью обучения с учителем является минимизация функционала риска $R(w)$ в классе функций аппроксимации $\{F(x, w), w \in W\}$. Однако оценка функционала риска усложняется тем, что обобщенная функция распределения $F_{x,d}(x, d)$ обычно неизвестна. При обучении с учителем вся доступная информация содержится в множестве данных обучения T . Чтобы обойти эту математическую сложность, будем использовать индуктивный принцип минимизации эмпирического риска. Этот принцип основан на доступности обучающего множества T , что идеально согласуется с философией нейронных сетей [1].

В нашей ситуации общее понятие риска будет трактоваться следующим образом [2].

Определение 1. Совокупным риском развития неблагоприятной ситуации в сельском хозяйстве в области A назовем зависимость

$$R_A = \iint_A P(x, y)h(x, y)dx dy, \quad (1)$$

где $P(x, y)$ — вероятность неблагоприятного события в точке (x, y) , определяемая на основе совместного анализа разнородной информации, $h(x, y)$ — функция ущерба в позиции (x, y) , связанного с этим событием. Способ вычисления вероятности $P(x, y)$ будет описан ниже, а функции ущерба $h(x, y)$ определяется стоимостью сельскохозяйственной продукции в точке (x, y) .

Вероятность стихийного бедствия $P(x, y)$ определяется различными факторами окружающей среды и метеорологическими условиями, которые могут быть прямо или косвенно измерены с помощью наземных средств и дистанционных методов, либо получены с помощью моделирования. В качестве прогнозных моделей для оценки вероятности стихийных бедствий можно использовать различные регрессионные модели или модели в виде черного ящика (например, нейросетевые, ядерные методы, методы статистической теории оценивания и т.п.).

Таким образом, для оценки состояния сельскохозяйственных культур и вероятности риска неблагоприятного события, например, засухи, необходимо анализировать (классифицировать) информацию, поступающую из различных источников с различным временным и пространственным разрешением. Для совместного анализа такой информации применяются методы и технологии слияния данных (data fusion) [2, 3]. Для вычисления вероятности неблагоприятного события $P(x, y)$ предлагается следующая общая схема (рис. 1).

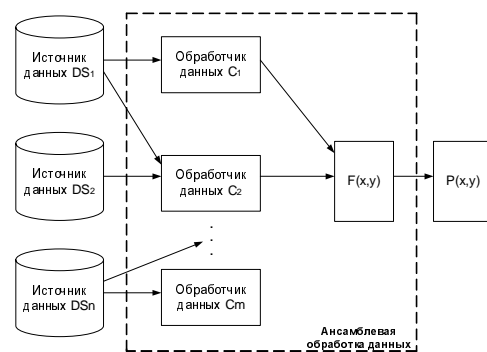


Рисунок 1— Схема определения вероятности неблагоприятного события $P(x, y)$ в точке (x, y) на основе разнородной информации

На рис. 1 блоки $DS_i, i=1..n$ представляют разные источники данных — спутниковые данные, данные наземных сенсоров и данные моделирования. Блоки $C_i, i=1..m$ обеспечивают обработку данных и получение информации более высокого уровня. В общем случае число блоков обработки не равно количеству источников данных n ($m \neq n$), поскольку данные из одного

источника могут обрабатываться разными обработчиками и, в то же время, один обработчик может анализировать данные из разных источников одновременно. То есть слияние данных происходит уже на стадии их преобразования данных обработчиками. Вид преобразования в каждом из блоков определяется C_i , $i=1\dots m$ типом входных данных. Например, спутниковые данные проходят несколько уровней предварительной обработки (перепроектирование, геопривязка, коррекция и т.д.), а затем подлежат так называемой «тематической» обработке. Будем иметь в виду, что каждый из обработчиков C_i , $i=1\dots m$ может обеспечивать несколько уровней преобразования входных данных, но для ясности при дальнейшем изложении явно не будем выделять эти уровни.

Без ограничения общности можно считать, что каждый из обработчиков C_i , $i=1\dots m$ представляет собой отдельное решающее правило или классификатор (так называемый «слабый» или компонентный классификатор), обеспечивающий результат анализа данных одного или нескольких источников. Таким образом, классификаторы C_i , $i=1\dots m$ составляют ансамбль экспертов (или «сильный» классификатор), «мнения» которых с соответствующими весами α_i учитываются в блоке слияния F

$$F(x, y) = \sum_{i=1}^m \alpha_i C_i(x, y). \quad (2)$$

Представленная на рис. 1 схема обработки данных реализует ансамблевый метод анализа данных [4], обеспечивающий более высокую точность оценивания разнородной информации, чем любой отдельно взятый классификатор. Отдельные классификаторы выступают в такой схеме в качестве «ядер», обеспечивающих переход от анализа отдельных признаков к оценке ситуации в целом. Преимуществом ансамблевого метода классификации является возможность повышения точности классификации за счет приема «усиления» (boosting) [5], который состоит в оценке функции потерь (ошибки классификации) и минимизации этой функции путем добавления новых компонентных классификаторов. Если добавление нового «слабого» классификатора не снижает значение функции потерь, этот классификатор не привносит новых информативных признаков для оценки ситуации и его можно не учитывать. Таким образом, обеспечивается оптимальная сложность модели [6].

Минимизация функции потерь на всем множестве данных в теории машинного обучения называется задачей минимизации среднего риска, а задача минимизации ошибки классификации на обучающем множестве называется минимизацией

эмпирического риска [1].

Таким образом, оценка рисков стихийных бедствий на основе разнородной информации сводится к задаче минимизации эмпирического и среднего риска при обучении классификатора. Рассмотрим математический аппарат минимизации рисков при обучении.

Математический аппарат минимизации рисков при обучении классификатора

Основная идея принципа минимизации эмпирического риска (empirical risk minimization) состоит в использовании функционала эмпирического риска $R_{emp}(w)$ [1], определяемого формулой.

$$R_{emp}(w) = \frac{1}{N} \sum_{i=1}^N L(d_i, F(x_i, w)). \quad (3)$$

Пусть w_{emp} и $F(x, w_{emp})$ - вектор весов и соответствующее ему отображение, которые минимизируют функционал эмпирического риска $R_{emp}(w)$. Аналогично, пусть w_0 и $F(x, w_0)$ - вектор весовых коэффициентов, и отображение, минимизирующие фактический функционал риска $R(w)$. Векторы w_{emp} и w_0 принадлежат пространству весов W . Требуется найти условия, при которых аппроксимирующее отображение $F(x, w_{emp})$ достаточно "близко" к фактическому отображению $F(x, w_0)$ (в качестве меры близости будем использовать разницу между $R_{emp}(w)$ и $R(w)$).

Для некоторого фиксированного $w = w^*$ функционал риска $R(w^*)$ определяет математическое ожидание случайной переменной, определяемое соотношением

$$Z_{w^*} = L(d, F(x, w^*)) \quad (4)$$

В отличие от него функционал эмпирического риска $R_{emp}(w^*)$ обеспечивает эмпирическое (арифметическое) среднее значение (empirical (arithmetic) mean) случайной переменной Z_{w^*} . Согласно закону больших чисел (law of large numbers), который составляет одну из основных теорем теории вероятностей, для обучающего множества T бесконечно большого размера N эмпирическое среднее случайной переменной Z_{w^*} в общем случае сходится к ее ожидаемому значению. Это наблюдение обеспечивает теоретический базис для использования функционала эмпирического риска $R_{emp}(w)$ вместо функционала риска $R(w)$.

Однако, тот факт, что эмпирическое среднее переменной Z_w , сходится к ее ожидаемому значению, совершенно не означает, что вектор весовых коэффициентов w_{emp} , минимизирующий функционал эмпирического риска $R_{emp}(w)$, будет также минимизировать и функционал риска $R(w)$.

Этому требованию можно приближенно удовлетворить, применив следующий подход. Если функционал эмпирического риска $R_{emp}(w)$ аппроксимирует исходный функционал риска $R(w)$ равномерно по w с некоторой точностью ε , то минимум $R_{emp}(w)$ отстоит от минимума $R(w)$ не более чем на величину 2ε . Формально это означает необходимость обязательного выполнения следующего условия. Для любого $w \in W$ и $\varepsilon > 0$ должно выполняться вероятностное соотношение

$$P(\sup_w |R(w) - R_{emp}(w)| > \varepsilon) \rightarrow 0 \text{ при } N \rightarrow \infty \quad (5)$$

Если выполняется условие (5), то можно утверждать, что вектор весов w среднего эмпирического риска равномерно сходится к своему ожидаемому значению. Таким образом, если для любой наперед заданной точности ε и некоторого положительного α выполняется неравенство

$$P(\sup_w |R(w) - R_{emp}(w)| > \varepsilon) < \alpha \quad (6)$$

то выполняется также и следующее неравенство:

$$P(R(w_{emp}) - R(w_0) > 2\varepsilon) < \alpha \quad (7)$$

Другими словами, если выполняется условие (6), то с вероятностью $(1-\alpha)$ решение $F(x, w_{emp})$, минимизирующее функционал эмпирического риска $R_{emp}(w)$, обеспечивает отличие фактического риска $R(w_{emp})$ от минимально возможного фактического риска на величину, не превышающую 2ε . Это значит, что при выполнении (6) с вероятностью $(1-\alpha)$ одновременно выполняются следующие два неравенства:

$$R(w_{emp}) - R_{emp}(w_{emp}) < \varepsilon \quad (8)$$

$$R_{emp}(w_0) - R(w_0) < \varepsilon \quad (9)$$

Эти два соотношения определяют различие между функционалами истинного и эмпирического рисков в точках $w = w_{emp}$ и $w = w_0$. Учитывая, что w_{emp} и w_0 являются

точками минимума функционалов $R_{emp}(w)$ и $R(w)$, можно сделать вывод о том, что

$$R_{emp}(w_{emp}) \leq R_{emp}(w_0). \quad (10)$$

Складывая неравенства (8) и (9) и принимая во внимание неравенство (10), можно записать:

$$R(w_{emp}) - R(w_0) < 2\varepsilon \quad (11)$$

Поскольку неравенства (8) и (9) одновременно выполняются с вероятностью $(1-\alpha)$, то с такой же вероятностью выполняется и неравенство (11). Также можно утверждать, что с вероятностью α будет выполняться неравенство

$$R(w_{emp}) - R(w_0) > 2\varepsilon \quad (12)$$

которое является еще одной формой утверждения (7).

Теперь можно формально сформулировать принцип минимизации эмпирического риска (principle of empirical risk minimization), состоящий из трех взаимосвязанных частей.

1. Вместо функционала риска $R(w)$ строится функционал эмпирического риска

$$R_{emp}(w) = \frac{1}{N} \sum_{i=1}^N L(d_i, F(x_i, w))$$

на базе множества примеров обучения (x_i, d_i) , $i = 1, 2, \dots, N$.

2. Пусть w_{emp} - вектор весовых коэффициентов, минимизирующий функционал эмпирического риска $R_{emp}(w)$ в пространстве весов W . Тогда $R_{emp}(w)$ сходится по вероятности к минимально возможным значениям фактического риска $R(w)$, $w \in W$. При этом при увеличении количества N примеров обучения до бесконечности функционал эмпирического риска $R_{emp}(w)$ равномерно сходится к функционалу фактического риска $R(w)$.

3. Равномерная сходимости, определяемая как

$$P(\sup_{w \in W} |R(w) - R_{emp}(w)| > \varepsilon) \rightarrow 0 \text{ при } N \rightarrow \infty$$

является необходимым и достаточным условием непротиворечивости принципа минимизации эмпирического риска.

Для физической интерпретации этого важного принципа проведем следующее наблюдение. До обучения машины все аппроксимирующие функции равноценны. По мере обучения правдоподобие тех функций аппроксимации, которые не противоречат обучающему множеству $\{(x_i, d_i)\}_{i=1}^N$ возрастает. По мере увеличения количества использованных при обучении примеров и, следовательно, повышения "плотности" входного пространства точка минимума функционала эмпирического

риска $R_{emp}(w)$ сходиться по вероятности к точке минимума функционала фактического риска $R(w)$.

Оценка риска неблагоприятных ситуаций в сельском хозяйстве

Для оценивания риска неблагоприятных ситуаций в сельском хозяйстве использовались разные нейросетевые методы классификации спутниковых изображений: классификация с помощью многослойного перцептрона (MLP), гауссовского классификатора максимального правдоподобия (ML) и нейронной сети типа ARTMAP.

Классификация с помощью многослойного перцептрона (MLP). При классификации с использованием многослойного перцептрона (MLP) [1] процедура перекрестной проверки выполнялась для нейронной сети с 1 скрытым слоем и разным количеством скрытых нейронов: 5, 15, 20, 25, 35 и 45. Для обучения нейронных сетей использовался метод EBD [7]. Количество эпох составляло 500. Начальное значение коэффициента обучения равнялось 0,7, а значение фактора момента – 0,5. Наибольшее значение числа верно классифицированных векторов было получено для нейронной сети с 45 скрытыми нейронами. (табл. 1)

Табл.1 Статистика классификации для обучающего множества по результатам работы многослойного перцептрона

Класс	Архитектура нейронной сети (количество скрытых нейронов)			
	5	25	35	45
1	97.63	99.02	99.15	98.97
2	80.95	84.20	84.64	85.67
3	67.09	68.38	68.00	67.37
4	85.44	89.03	89.84	89.56
5	86.16	90.41	91.01	91.43
6	97.14	97.75	97.63	97.64
7	69.09	83.99	83.46	83.56
8	95.57	96.53	96.79	96.52
Всего	84.88	88.68	88.81	88.85

Гауссовский классификатор максимального правдоподобия. Матрицы математического ожидания и ковариации оценивались для каждого класса отдельно с использованием 5 обучающих последовательностей. Для этого метода использовались следующие стандартные оценки:

$$\hat{\mu}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} x_i^j,$$

$$\hat{\Sigma}_i = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (x_i^j - \hat{\mu}_i)(x_i^j - \hat{\mu}_i)^T,$$

где x_i^j - j-ая выборка i-го класса, M_i - номер

выборки в i-ом классе.

Результаты классификации ML сети для тестового множества представлены в таблице 2.

Таблица 2 Результаты работы MLP классификатора

Классы		
лиственный лес	1	98.73
хвойный лес	2	83.68
смешанный лес	3	67.68
пашня	4	89.66
пастбище	5	92.82
вода	6	96.57
искусственные материалы	7	82.18
открытая местность	8	96.75
Всего		84.88

Классификация с помощью ARTMAP.

Пятикратная перекрестная проверка проводилась для разных значений параметра ρ нейронной сети ARTMAP [8, 9] : 0.1, 0.2, 0.3, 0.5, и 0.95. Наибольшее значение числа верно классифицированных векторов было получено для нейронной сети со значением параметра подобия 0,95. Результаты классификации содержатся в таблице 3.

Таблица 3. Результаты работы нейронных сетей на основе теории адаптивного резонанса

Класс	Параметры подобия				
	0,1	0,2	0,3	0,5	0,95
1	98,92	99,68	99,56	98,52	99,88
2	79,58	80,86	80,34	79,16	80,88
3	69,14	68,16	68,66	69,36	68,14
4	81,50	81,50	81,72	81,88	83,50
5	76,48	74,26	75,34	74,10	78,94
6	96,70	96,60	96,76	97,40	93,76
7	79,38	77,28	78,32	77,12	76,78
8	96,42	97,36	97,00	97,54	98,24
Всего	83,68	83,80	83,74	83,24	84,22

Сравнение результатов работы различных нейросетевых архитектур приводится в таблице 4.

Таблица 4. Сравнение результатов работы различных нейросетевых архитектур.

Класс	MLP	ML	ARTMAP
	98.97	98.73	99.88
1	85.67	83.68	80.88
2	67.37	67.68	68.14
3	89.56	89.66	83.50
4	91.43	92.82	78.94
5	97.64	96.57	93.76
6	83.56	82.18	76.78
7	96.52	96.75	98.24
8	88.85	88.02	84.22

Анализируя результаты проведенных экспериментов, можно утверждать, что для разных классов наилучший результат достигался разными методами. Это можно объяснить сложностью входных данных и сложными свойствами разделимости поверхностей классов в пространстве признаков. В среднем по всей выборке данных наилучшие результаты дает многослойный перцептрон, обученный методом обратного распространения ошибки. Для класса 2,6 и 7 MLP классификатор превосходит ARTMAP и ML классификаторы. В свою очередь, нейронная сеть ARTMAP была лучше для классификации классов 1, 3, 8, и ML классификатор лучше определял классы 4 и 5.

Хуже всего классификаторы определяли класс 3 «смешанные леса» (ARTMAP - 68.14%). Это связано с тем, что смешанные леса (класс 3) состоят из лиственных (класс 1) и хвойных лесов (2), что сказывается на спектральных характеристиках.

На рис. 2 представлена маска посевов озимой пшеницы за 2008 год, полученная по данным спутника LandSAT.



Рисунок 2 – Маска площадей озимых культур для территории Украины

Ансамблевая нейронная сеть

Для повышения эффективности классификации и интегрированной оценки рисков неблагоприятных ситуаций в сельском хозяйстве предлагается использовать ансамблевую обработку данных, схема которой представлена на рис. 1. В качестве отдельных обработчиков данных можно применять различные классификаторы, в том числе ансамблевые нейронные сети. Нейронная сеть называется ансамблевой, если ее можно разбить на две или более подсистемы, которые функционируют на определенных входных данных без взаимодействия между собой.

Преимущество использования ансамблевых нейронных сетей:

– Возможность разделить задачу на несколько подзадач и определить более подходящий алгоритм решения подзадач;

– Снижение сложности архитектуры нейронной сети

– Уменьшение времени обучения.

Для решения поставленной задачи предлагается использовать следующий ансамблевый метод обработки данных:

– В качестве классификаторов (обработчиков данных) выбираются различные архитектуры нейронных сетей, в том числе многослойный перцептрон MLP (первый столбец в табл.5)

– Применяется метод: «один класс» - «одна нейросеть». Для каждого класса строится своя нейросеть, которая должна отличать этот класс от других (One-One) (второй столбец в табл. 5).

– Блок слияния Arch (модуль F) представляет собой многослойный перцептрон с алгоритмом EDBD (Enhanced Delta-Bar-Delta) [10] (третий столбец в табл. 5).

Сравнительные результаты ансамблевой обработки данных приводятся в табл. 5.

Таблица 5. Сравнение результатов работы различных методов классификации (жирным выделен наилучший результат для данного класса)

Класс	Метод		
	MLP (6-45-8)	One-One	Arch
1	98.97	99.39	99.19
2	85.67	84.44	84.24
3	67.37	67.51	68.64
4	89.56	89.18	89.61
5	91.43	92.48	93.08
6	97.64	97.15	97.97
7	83.56	79.63	84.20
8	96.52	97.86	96.79
Всего	88.85	88.45	89.21

На рис. 3 и 4 изображены реализованные сервисы определения биоразнообразия и оценки рисков засух в Украине, которые базируются на описанных выше методах ансамблевой обработки данных с учетом минимизации функционала эмпирического риска при обучении.

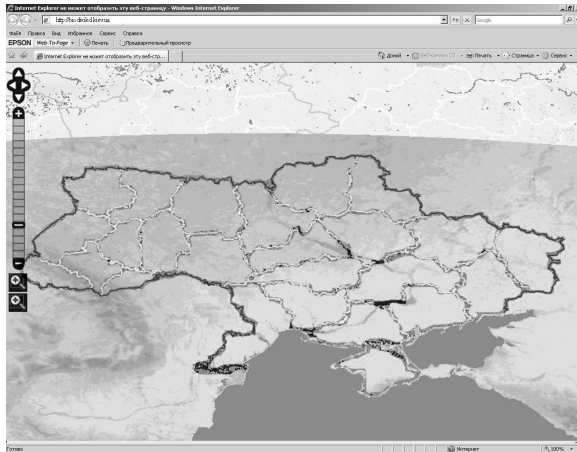


Рисунок 3 – Сервис определения биоразнообразия

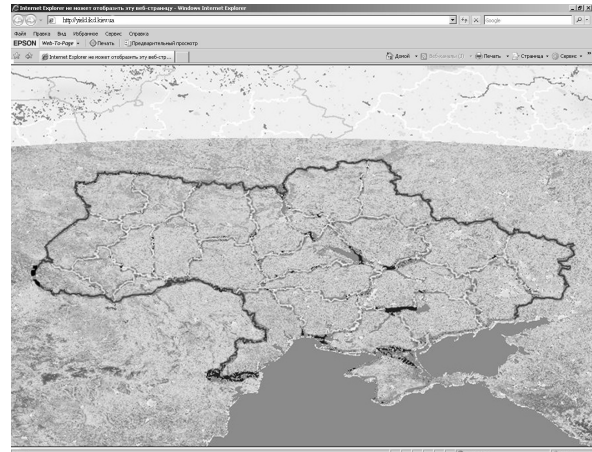


Рисунок 4 – Сервис оценки риска засух по
разнородным данным

Выводы

В данной статье предлагается постановка задачи оценки рисков развития сельскохозяйственных культур и обосновывается применяемый для решения этой задачи математический аппарат, а также приводятся результаты его практического применения для

оценки рисков засух и оценки биоразнообразия в Украине. Показано, что ансамблевая обработка данных повышает эффективность решения задачи классификации состояния растительности и оценки рисков развития сельскохозяйственных культур.

Литература

1. Хайкин С. Нейронные сети: полный курс, 2-е издание.: Пер. с англ. - М. Издательский дом «Вильямс», 2006. - 1104 с.
2. Н.М. Куссуль, Я.И. Зельк, Интегрированный подход к оценке рисков на основе геопространственной информации// Труды Втроей Украинской конференции с международным участием «Аэрокосмические наблюдения в интересах устойчивого развития и безопасности». — К.: Наукова думка, 2010. — С. 16-17.
3. S. Das, High-Level Data Fusion (2008), Artech House Publishers, Norwood, MA, ISBN 9781596932814 and 1596932813
4. Tommi Jaakkola, course materials for 6.867 Machine Learning, Fall 2006. MIT OpenCourseWare(<http://ocw.mit.edu/>), Massachusetts Institute of Technology.
5. S. Kotsiantis, P. Pintelas, Combining Bagging and Boosting, International Journal of Computational Intelligence, Vol. 1, No. 4 (324-333), 2004.
6. M. Piers, Methods and models for the assessment of third party risk due to aircraft accidents in the vicinity of airports and their implications for societal risk// In: R.E. Jorissen, P.J.M. Stallen (Eds.), Quantified Societal Risk and Policy Making, Kluwer Academic Publishers, Dordrecht, 1998.
7. P.J. Werbos. The roots of backpropagation: from ordered derivatives to neural networks and political forecasting, John Wiley & Sons, Inc., New York, 1994, 319 p
8. G.A. Carpenter, S. Grossberg, J.H. Reynolds. ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network, Neural Networks Vol. 4 (1991). pp. 565-588.
9. G.A. Carpenter, S. Grossberg. ART 2: Stable selforganization of pattern recognition codes for analog input patterns, Applied Optics, vol. 26 (1987). pp. 4919-4930.
10. A.A. Minai, R.J. Williams. Back-propagation heuristics: A study of the extended delta-bar-delta algorithm. IEEE International Joint Conference on Neural Networks vol. I (1990). pp. 595-600

Поступила в редакцию 30.03.2010